



## King's Research Portal

DOI:

[10.1002/hec.4025](https://doi.org/10.1002/hec.4025)

*Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Carrino, L., Glaser, K., & Avendano Pabon, M. (2020). Later retirement, job strain, and health: Evidence from the new State Pension age in the United Kingdom. *Health Economics*, 29(8), 891-912.  
<https://doi.org/10.1002/hec.4025>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

RESEARCH ARTICLE

Health  
Economics

WILEY

# Later retirement, job strain, and health: Evidence from the new State Pension age in the United Kingdom

Ludovico Carrino<sup>1,2</sup>  | Karen Glaser<sup>1</sup>  | Mauricio Avendano<sup>1,3</sup> 

<sup>1</sup>Department of Global Health & Social Medicine, King's College, London, UK

<sup>2</sup>Department of Economics, University of Venice Ca' Foscari, Venice, Italy

<sup>3</sup>Department of Social and Behavioural Sciences, Harvard School of Public Health, Boston, Massachusetts, USA

## Correspondence

Ludovico Carrino, Department of Global Health & Social Medicine, King's College London, Bush House North East Wing, 40 Aldwych London (WC2B4BG), UK.  
Email: ludovico.carrino@kcl.ac.uk

## Funding information

Economic and Social Research Council, Grant/Award Numbers: ES/P000487/1 (Worklong, Joint Programme Initiative 'More Years, Better Lives'), ES/LS002825/1 (Wherl), ES/S012567/1 (ESRC Centre for Society and Mental Health at King's College London); European Commission Horizon 2020 Programme, Grant/Award Number: 667661 (Mindmap)

## Abstract

This paper examines the impact of raising the State Pension age on women's health. Exploiting a UK pension reform that increased women's State Pension age for up to 6 years since 2010, we show that raising the State Pension age leads to an increase of up to 12 percentage points in the probability of depressive symptoms, alongside an increase in self-reported medically diagnosed depression among women in a lower occupational grade. Our results suggest that these effects are driven by prolonged exposure to high-strain jobs characterised by high demands and low control. Effects are consistent across multiple subcomponents of the General Health Question and Short-Form-12 (SF-12) scores, and robust to alternative empirical specifications, including "placebo" analyses for women who never worked and for men.

## KEYWORDS

economics of ageing, public health, public pensions, social security, understanding society

## JEL CLASSIFICATION

J26; H75; I18; I14

## 1 | INTRODUCTION

Over the last decade, most Organisation for Economic Cooperation and Development (OECD) countries have increased their statutory pensionable age (SPA) with the aim of enhancing the financial sustainability of pension systems (OECD, 2016). The rationale behind these reforms is that increased employment opportunities, longer life expectancy, and more years spent in good health will enable older people to work longer and retire later (OECD, 2016; OECD, 2017a). Although evidence suggests that health is the most important cause of early retirement (Munnell, Sanzenbacher, & Rutledge, 2015), understanding how policies delaying retirement influence health is critical to assessing the reforms' total utility (Heller-Sahlgren, 2017).

This study assesses the health impact of a recent reform that gradually increased the SPA from age 60 up to 66 years for women born after March 1950 in the United Kingdom. The effect of an increase in SPA is ambiguous as several mechanisms may be at play. The Grossman model considers health as both an investment good that increases productivity and a consumption good that provides utility (Galama & Kapteyn, 2011; Grossman, 1972). On the one hand,

-----  
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. Health Economics published by John Wiley & Sons Ltd

workers may invest more in their health if they expect to retire later, for example, by engaging in healthy behaviours if the benefits of a longer working life induced by better health are higher than the costs of reduced leisure-time due to a shorter retirement period (Bertoni, Brunello, & Mazzarella, 2018). Alternatively, as postponing retirement increases the opportunity cost of time, individuals might trade-off health investments (exercising, cooking healthier food, attending medical visits) for working time (Galama & Kapteyn, 2011). Later retirement may also directly enter the health production function, for example, through higher social engagement, mental wellbeing, cognitive function, and other non-financial benefits of work (Mazzonna & Peracchi, 2017). The type of occupation, however, may be critical to understanding the way these mechanisms operate. In particular, later retirement for workers in a low occupational grade may increase exposure to work-related psychological and physical strain, which may result in poorer health as a result of extended exposure to these factors (Galama & Kapteyn, 2011). This is consistent with recent literature suggesting that workers subject to high job strain, as measured by high job demand (physical and psychosocial) and low job control (limited decision authority and intellectual discretion), experience worse health outcomes than workers with occupations subject to low job strain (see Marmot et al. (1991), Karasek (1979), and Ravesteijn, Kippersluis, and Doorslaer (2018)).

Several studies examine how health is affected after retirement, with findings being sensitive to the choice of country, empirical strategy, and health outcome. Some studies find that retirement has positive effect on mental (Belloni, Meschi, & Pasini, 2016; Eibich, 2015) and physical health (Bertoni, Maggi, & Weber, 2017; Coe and Zamarro, 2011), while other studies report either a negative (Behncke, 2012; Bonsang, Adam, & Perelman, 2012; Mazzonna & Peracchi, 2017) or no effect of retirement on health (Coe and Lindeboom, 2008; Coe and Zamarro, 2011). By contrast, few studies have evaluated how recent reforms to the SPA influence the health of older people, and existing studies are mixed. A reform that postponed early retirement age by 5 years and reduced early pension replacement rates for civil servants in the Netherlands led to worse mental health (De Grip, Lindeboom, & Montizaan, 2012). In Israel, Shai (2018) found that a 2-year increase in the male SPA led to worsening health, as did Atalay and Barrett (2014) for an Australian reform that raised the female SPA by 5 years (spanned in 20 years). Bertoni et al. (2018) found that delaying statutory early retirement-age improved healthy behaviours and health satisfaction among Italian men in their 40s. Two studies evaluated the long-term impact of reforms lowering SPA, finding positive effects in the Netherlands within 5 years of retirement (Bloemen, Hochguertel, & Zweerink, 2017) and no effects in Norway (Hernaes, Markussen, Piggott, & Vestad, 2013).

Our study makes three important contributions to the literature. First, to our knowledge, this is the first study on the health effects of a unique UK reform that increased the SPA by up to 6 years over a short time window of 10 years. We differentiate from most earlier studies which examined how health changes after retirement as a result of cross-country variation in SPA (Mazzonna & Peracchi, 2017), or employer-based retirement windows (Behncke, 2012; Bonsang et al., 2012; Eibich, 2015), by focusing on a national policy change that affected the SPA of a well-defined cohort of women. We implement a difference-in-differences approach comparing the health status of women unable to collect their State Pension because of the SPA change with the health of women of similar age and characteristics who were unaffected by the change by virtue of their birthdate. Second, we are able to examine heterogeneity by job-type and job-stress in the causal effect of postponing SPA on health. Previous studies focusing on the health-transition after retirement found stronger effects on the health of workers from lower socioeconomic status (SES), who face lower life expectancy and income, more barriers to reemployment, and to good-quality care (Belloni et al., 2016; Bertoni et al., 2017; Coe, von Gaudecker, Lindeboom, & Maurer, 2012; Mazzonna & Peracchi, 2017). Yet, there is limited empirical evidence on whether increased exposure to job stressors might explain the impact of extended work horizon on health. By implementing validated indicators of job demand and job control, we investigate the role of longer exposure to job strain as a result of an increase in the SPA. Third, we are able to estimate nonlinearities in the effect of increases in SPA by comparing cohorts that experienced vastly different SPA extensions. This is possible because of the nature of the reform, which led to a relatively wide range of SPA increases (1 to 60 months) in a short period of a few years.

Using *Understanding Society*, a nationally representative survey with extensive health measures, we find that SPA increases had a negative impact on health: women aged 60–64 years who are no longer eligible to collect their pension due to the reform exhibit worse mental and physical health scores (PCSs) and higher prevalence of clinical depression than women of the same age unaffected by the reform. Moreover, longer extensions of SPA led to higher declines in mental health than shorter extensions. Crucially, the negative health effect of SPA postponement is confined to women from lower-grade routine occupations, and it is largely driven by longer exposure to adverse psychological and physical stressors. As a result, the reform had the undesirable consequence of increasing health inequality by occupational grade, as evidence points to a 12 percentage-point increase in the probability of depressive symptomatology (General

Health Question [GHQ] scores) for women in lower-grade occupations, which constitute clinically and economically meaningful changes. Moreover, we find a statistically significant 4.5% decline in PCSs for women in lower-grade occupations, although this is likely to be of less clinical relevance.

In what follows, Section 2 summarises the UK pension reform, Section 3 describes the data and empirical strategy, while Section 4 discusses the results. Section 5 concludes by discussing the implications of our findings.

## 2 | STATE PENSION AGE POSTPONEMENT IN THE UNITED KINGDOM

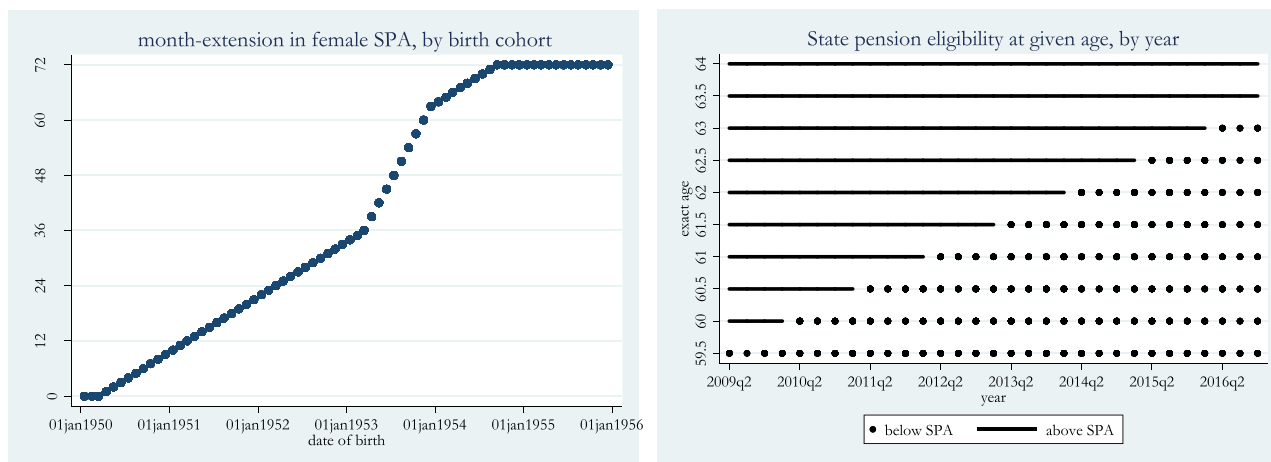
The reform of the female SPA in the United Kingdom, legislated in 1995 and implemented in 2010, affected the minimum age for claiming the Basic State Pension (BSP), which provides an almost-flat level of retirement income, depending on National Insurance contribution years. Although generally low by OECD standards, the BSP corresponds to around 38% of average gross income for retired households in the United Kingdom, thus representing a main income source for a significant proportion of older people (Webber & Mallet, 2017). In 2010, the full annual BSP amounted to £5,077.8 for a single individual and £8,119.8 for a couple (see OECD (2013), PPI (2015) and Lain (2016)). The SPA, which was set at 60 years prior to the reform, was first legislated to raise to 65 between 2010 and 2020 (an effective increase of 1 month every 2 months). Subsequent reforms then legislated for both men's and women's SPA to reach 66 by 2020, 67 by 2028, and 68 by 2046 (see Thurley and Keen (2017) for comprehensive details).

The impact of the reform on pension eligibility, based on birthdate, is substantial (Figure 1): being born 1 year after March 1950 implies a 1-year delay in SPA. The SPA postponement exceeds 36 months for cohorts born after March 1953. Thus, women born just a few years apart face different eligibility status (above vs. below SPA) at any given age (right panel): for example, a woman aged 60 years in 2009 is above her SPA, while a woman aged 60 years in 2012 is 2 years below her SPA. The Institute for Fiscal Studies estimated that the policy increased female employment rates by 10 percentage points (Cribb & Emmerson, 2018).

## 3 | METHODS

### 3.1 | Data

We use data from seven Waves (2009–2016) of *Understanding Society*, an annual survey interviewing household members aged 16+ years in Britain, on health, social, and economics subjects (Lynn (2009)). From information on year and month of birth and interview date, we determine whether an individual lies above or below her SPA when interviewed. We employ self-reported information on employment-status (“in-paid-work,” “unemployed,” “retired,”



**FIGURE 1** Change in women State Pension age.

Note. Authors' calculations based on Pension Act 1995, 2011 [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

“looking after family or home,” “long-term sick/disabled”), as well as information on living arrangements, employment history, number of children, educational attainment, and SES (detailed hereafter).

We employ three validated measures of mental and physical health (detailed in Appendix 1). The General Health Questionnaire index (GHQ-12) measures psychological distress, on a scale between 0 and 36, with higher values signalling worse health. As a score of 12+ signals the presence of common mental disorders (Goldberg et al., 1997; Goldberg & Williams, 1988), a GHQ-caseness is built to identify respondents lying above and below the cutoff. We further disaggregate the GHQ score in three clinically meaningful factors (anxiety/depression, social-dysfunction, and loss-of-confidence), following Graetz (1991). We normalise the GHQ scores between 0 and 100.

The Short-Form-12 (SF-12, version2) is a generic health-related quality of life instrument which produces a PCS and a mental health score (MCS), each ranging from 0 to 100, with higher values signalling better health. Eight meaningful subfactors are also considered: physical functioning, role-limitations due to (a) physical or (b) emotional problems, mental health, bodily pain, general health, vitality/energy/fatigue, and social functioning (methodological details are available in Ware (2002) and Appendix 1). Both the GHQ and the SF-12 are widely used in the economics literature as generic measures of physical and mental health (Bünnings, Kleibrink, & Weßling, 2017; Clark, 2003; Dustmann & Fasani, 2016; Marcus, 2013; Mitra & Jones, 2017; Schmitz, 2011). In addition, we also employ information on chronic diseases that respondents report to have been diagnosed with by a doctor.<sup>1</sup>

### 3.2 | Meaningful changes and effect size

In order to evaluate the clinical relevance of any impact of the SPA reform on health, we follow two approaches. First, we compare our results for continuous health outcomes to the concept of the Minimally Important Difference (MID), that is, a change large enough to be discernible by patients. The general approach to MID is to compute the “effect-size” (ES), that is, the ratio between the estimated effect and the standard deviation of that outcome (Cohen, 2013). Fayers and Hays (2014) suggest that an ES of 0.2 *SD* represents the threshold for a MID (see also Farivar, Liu, and Hays (2004)). Cohen (2013) proposed operational definitions of small, medium, and large ES corresponding to 0.2, 0.5, and 0.8 *SD*, respectively (see also King (2011)). For the PCS outcome, we will also adopt the threshold suggested by Schmitt and Di Fabio (2004), who estimated an “anchoring” minimum change in the PCS (6.8 points) which was found to correspond to a unitary change in patients’ Global Disability Rating.<sup>2</sup> Indeed, in lack of an accepted gold-standard, and given that both effects-size and anchoring are not immune to criticisms, the clinical literature recommends, when feasible, to use both methods in order to evaluate MIDs (Fayers & Hays, 2014; Jayadevappa, Cook, & Chhatre, 2017).

Second, we estimate the impact of the reform on the aforementioned GHQ-caseness index.

### 3.3 | Econometric specification

We examine the impact of raising the SPA on women’s health using a difference-in-differences approach on a narrow sample of women aged 60 to 64 years between 2009 and 2016. We compare the health status of respondents who, at the time of interview, were ineligible for the State Pension because of the reform with that of women of similar age who were eligible to claim their State Pension and were never affected by the reform (see Shai (2018) and Bertoni et al. (2018)).<sup>3</sup> Age groups 60–63 years are observed both below and above-SPA, while 64 years old are above-SPA in all years (Figure 1, right panel). Our identification strategy exploits variation in exposure to the reform by birth and interview dates and relies on a comparison of health trends for age cohorts affected by the change (the 60, 61, 62, and 63-year-olds born after April 1950 became ineligible after 2010, 2012, 2014, and 2016, respectively) relative to health trends in age cohorts never affected by the reform (60–64-year-olds born before April 1950 are always eligible to claim their State Pension).

<sup>1</sup>Self-reported diagnosis may be affected by measurement error, as some diseases remain undiagnosed, potentially leading to a zero-effect bias.

<sup>2</sup>The physical health score (PCS) anchoring suggested by Schmitt and Di Fabio (2004) is based on a sample of 155 patients with upper extremity diagnoses in Minneapolis, Minnesota.

<sup>3</sup>To ensure a “clean” control-group, we exclude respondents above their statutory pensionable age (SPA) who had experienced an earlier postponement in their SPA (i.e., previously affected by the reform). Results do not change in analyses that include these respondents (Appendix 1.5).



Starting from an eligible sample of 13,084 observations (4,925 women) aged 60 to 64 years between 2009 and 2016, we dropped 2,004 observations (15.3% of the eligible sample) of women who never worked (not affected by the SPA increase), proxy interviews (235 observations, 1.8%), entries with missing values in our health outcomes or in other control variables (1,162 observations, 8.9%), and 2,309 observations of women observed past their State Pension age but who had been previously affected by the reform (17.6%). Our working-sample thus comprises 7,374 observations (3,531 women).

We estimate the following reduced-form model for health outcome  $y_{iat}$ , observed at time  $t$  for individual  $i$  of age  $a$ , born in year-month  $c$ :

$$y_{iat} = \alpha + \beta \cdot \text{belowSPA}_{iat} + \gamma_a + \eta_t + \delta_c + X'_{iat}\varphi + \varepsilon_{iat}, \quad (3.1)$$

where our main independent variable of interest is an indicator function for being below SPA, that is, an interaction between the individual's age and the interview date, which captures whether individuals were eligible to claim a State Pension. As age is an important determinant of both State Pension eligibility and vulnerability to health shocks, we incorporate fixed effects for age quarters ( $\gamma_a$ ) to control for age-specific effects. Similarly, we include fixed effects for interview year quarters ( $\eta_t$ ) to capture common trend shocks in health outcomes as well as in employment rates. Finally, we include a linear control for year-month of birth ( $\delta_c$ ) to capture cohort effects in health and labour market attachment. These three variables are not collinear as they are included in different units and functional forms; in addition, we observe women of the same age born in different years and measured at different times. As noted by, for example, Cribb, Emmerson, and Tetlow (2016), such a model assumes that age effects are cohort- and time-constant, cohort effects are time- and age-constant, and time-effects are age- and cohort-constant, where the latter hypothesis represents a “common-trends assumption” which we will discuss in Section 4.2.4, alongside further tests for alternative parametric forms for age-time-cohort effects.

We control for additional observable individual- or area-specific characteristics that might confound the analysis (Staubli & Zweimüller, 2013). Due to cohort effects, treatment and control groups might differ on several sociodemographic characteristics. As younger cohorts in the United Kingdom have lower fertility (Kneale & Joshi, 2008), and having children could affect both women's employment/retirement decision and the onset of health conditions (Behncke, 2012), we include a categorical variable for having zero, one/two, or at least three children. Similarly, as younger cohorts are also more likely to attain secondary education and be employed in intermediate/managerial occupations than older cohorts (Mazzonna & Peracchi, 2017; OECD, 2015), we add one categorical variable for highest educational attainment (A-level or higher, GCSE-level, less than GCSE) and one for having (having had) a routine-, intermediate-, or managerial-level job (NS-SEC classification).<sup>4</sup> We account for marital status (being single, widowed/separated, or living with someone) to account for cohort differences in family forms and living arrangements (Sobotka & Toulemon, 2008). We control for country dummies (within the United Kingdom) to account for geographical factors that might directly affect health. Conditional on these controls, our coefficient  $\beta$  captures the impact of being below the SPA as a result of the reform, above and beyond the effect of age, year, and cohort.

We test for health inequality effects of postponing the SPA, by estimating a model including an interaction term between the “policy variable” and the NS-SEC occupational classification (or, in alternative models, the level of job-strain) (3.2) as follows:

$$y_{iat} = \alpha + \beta_1 \cdot \text{belowSPA}_{iat} + \beta_2 \cdot \text{intermediate}_{iat} + \beta_3 \cdot \text{manager}_{iat} + \beta_4 \cdot \text{belowSPA}_{iat} \cdot \text{intermediate}_{iat} + \beta_5 \cdot \text{belowSPA}_{iat} \cdot \text{manager}_{iat} + \gamma_a + \eta_t + \delta_c + X'_{iat}\varphi + \varepsilon_{iat}. \quad (3.2)$$

In all analyses, standard errors are clustered at the month-of-birth level (154 clusters), as the treatment assignment varies by month-of-birth. Findings are robust to individual-level clustering (Section 4.2.4).

<sup>4</sup>The NS-SEC class is based on current/last occupational class, firm-size, and employment type (employer/self-employed/employee)

### 3.3.1 | Effect of different levels of SPA postponement

We test how the health impact of the reform differs with the extent of SPA postponement, which widely differs for women born after March 1950. We modify specification (3.1) as follows:

$$y_{iat} = k + \lambda_1 I(1 \leq \text{months} \leq 6)_{ia} + \lambda_2 I(6 < \text{months} \leq 24)_{ia} + \lambda_3 I(24 < \text{months} \leq 36)_{ia} + \lambda_4 I(36 < \text{months})_{ia} + \mu_a + \alpha_t + \nu_c + X'_{iat} \theta + v_{iat}, \quad (3.3)$$

where we introduce dummies for having an SPA increase of 1–6, 7–25, 25–36,  $\geq 36$  months (the reference category is “no SPA-postponement”). The SPA increase is nonlinearly related to birthdate, as the SPA is constant for women born before April 1950 or after September 1954 while it increases nonlinearly depending on month-of-birth for women born between March 1950 and September 1954 (Figure 1, left panel).

## 3.4 | Descriptive evidence

Column (i) of Table 1 summarises descriptive statistics for our original eligible sample, while Column (ii) describes the final sample used in the statistical analysis. Around 40% falls into the manual-routine SES (mostly personal service occupations, sales and customer services, process/plant/machine operatives, and elementary occupations); around 30% belongs to the intermediate SES (mostly administrative and secretarial positions), and 30% falls into the higher SES group (managers and senior officials, health, teaching, and science professionals).<sup>5</sup> As Columns (iii) to (v) show, women ineligible for pension exhibit higher employment, unemployment, and sick/home-carers rates, as-well-as worse MCS and GHQ scores. Appendix Table 10 (Data S1) details the sample's age decomposition.

This pattern is further illustrated in Figure 2, where we compare employment and health outcomes for a control group always observed above-SPA (women aged 62–64 years) and two treatment groups whose SPA-status changes over the study period as a result of the reform (60–61 years old). Panels a and b show that retirement rates decrease while working/being sick/caring rates increase for the treated groups when they become ineligible for State Pension, while trends are stable for the control group (see also Cribb et al. (2016) and Staubli and Zweimüller (2013)). Interestingly, Panels d and e show that women observed below-SPA fare worse GHQ and MCS scores. These trends, however, could reflect cohort effects or omitted variable bias; in the next section, therefore, we turn to our econometric approach to isolate the casual impact of the SPA reform.

## 4 | RESULTS

### 4.1 | Health effects of the SPA reform

Table 2 reports the results for our main specification (3.1), estimated through OLS for the continuous health outcomes (GHQ, MCS, PCS) and through Linear Probability Model for the GHQ cutoff. In Column 1, we estimate the effect of the reform on female employment rates (measured with a binary indicator for being in paid work). We confirm previous findings (Cribb et al., 2016; Cribb & Emmerson, 2017) of a major shock to the employment of women whose SPA was postponed, estimated in a 10-percentage point increase. Self-reported alternative labour market outcomes are also affected, that is, respondents declaring to be sick (+5%), caring for the house/for someone (+2.5%), unemployed (+2%), and retired (−21%). Table 11 includes full results.<sup>6</sup>

<sup>5</sup>Compared with our original eligible sample, the final sample is very similar in terms of average age, marital status, and number of children, yet it exhibits generally higher education level and employment-rates, lower retirement/inactive rates, and slighter better average health scores. These differences are likely due to the exclusion of women who never worked, who are more likely to be low-educated, retired or inactive, and in worse health.

<sup>6</sup>Running models for “being in paid work” with interaction terms between “being below SPA” and education, marital status, number of children, and job-category, we find no statistically significant heterogeneous effects (available upon request), suggesting that the impact of the reform was quite similar across social groups.

**TABLE 1** Descriptive statistics for the whole sample and by pension eligibility status at interview date

	(i) Original eligible sample Mean	(ii) Final sample Mean	(iii) Above-SPA at interview Mean	(iv) Below-SPA at interview Mean	(v) <i>p</i> value (iii) v (iv) (controlled for birth-year)
Age	62.1	62.4	63.1	61.1	***
Married/couple	70.4%	70.6%	70.9%	70%	
Widowed/divorced	24.7%	24.6%	25%	23.9%	
Single	4.9%	4.7%	4%	6.2%	
Number of children	1.7	1.8	1.9	1.7	
Educ.: less than GCSE	23.9%	19.9%	23.9%	13%	
Educ: GCSE	33.8%	34.6%	35.2%	33.6%	
Educ: A-level or more	42.2%	45.5%	41%	53.8%	
SPA postponement (months)		15.7	0	43	***
Occupational status					
Employed	35.7%	43.5%	35.5%	57.6%	**
Unemployed	1.4%	1.3%	0.3%	3.1%	*
Retired	53.5%	47.9%	59.7%	26.7%	***
Sick/disabled	4.1%	3.3%	1.5%	6.7%	***
Caring for family/home	5.3%	3.7%	2.8%	5.1%	
Socioeconomic class (NS-SEC3)					
Routine	39.7%	38.6%	41.4%	33.3%	
Intermediate	27.6%	28.1%	28.6%	26.9%	
Higher occupations	32.7%	33.3	30.4%	38.4%	
Health measures					
SF-12 PCS (physical)	46.2 <sup>a</sup>	46.8	46.7	47.1	
SF-12 MCS (mental)	50.6 <sup>a</sup>	51	51.5	49.9	
GHQ-12 (0–100 scale)	30.9 <sup>b</sup>	30.4	29.8	31.4	**
GHQ cutoff	36% <sup>b</sup>	33%	33%	37%	*
With long term conditions at baseline					
Arthritis	31.3%	31.1%	33.9%	25.6%	
Coronary heart disease	1.1%	1.2%	1.4%	0.7%	
Angina	2.8%	2.8%	3.4%	1.8%	
Heart attack	1.9%	1.9%	2.1%	1.4%	
Liver condition	1.1%	1.2%	1.2%	1.3%	
Diabetes	7.5%	6.9%	8.2%	4.4%	
High blood pressure	25.1%	25.1%	28%	20%	
Clinical depression	6.5%	5.9%	5%	7.5%	
Observations	13,084	7,374	4,740	2,634	

*Note.* The final sample includes women aged 60–64 between 2009 and 2016, observed either above-SPA (never affected by the reform) or below-SPA (affected by the reform), having been engaged in paid work in their life. Column (iv) reports the test for the null-hypothesis of mean-equivalence between Columns (ii) and (iii). The status of being above/below SPA is defined by comparing the individual SPA (based on month-year of birth) and the date of interview. The SPA postponement is a distance measured in months between the individual-specific SPA postreform and the prereform threshold of 60 years old. The job classification follows the National Statistics SEC-3 taxonomy. The GHQ cut-off refers to the Likert GHQ scale (range: 0–36) and takes value 1 for scores of 12+ (Goldberg et al., 1997).

Abbreviations: GHQ, General Health Question; MCS, mental health score; PCS, physical health score; SF-12, Short-Form-12; SPA, State Pension age.

<sup>a</sup>*N* = 11,514.

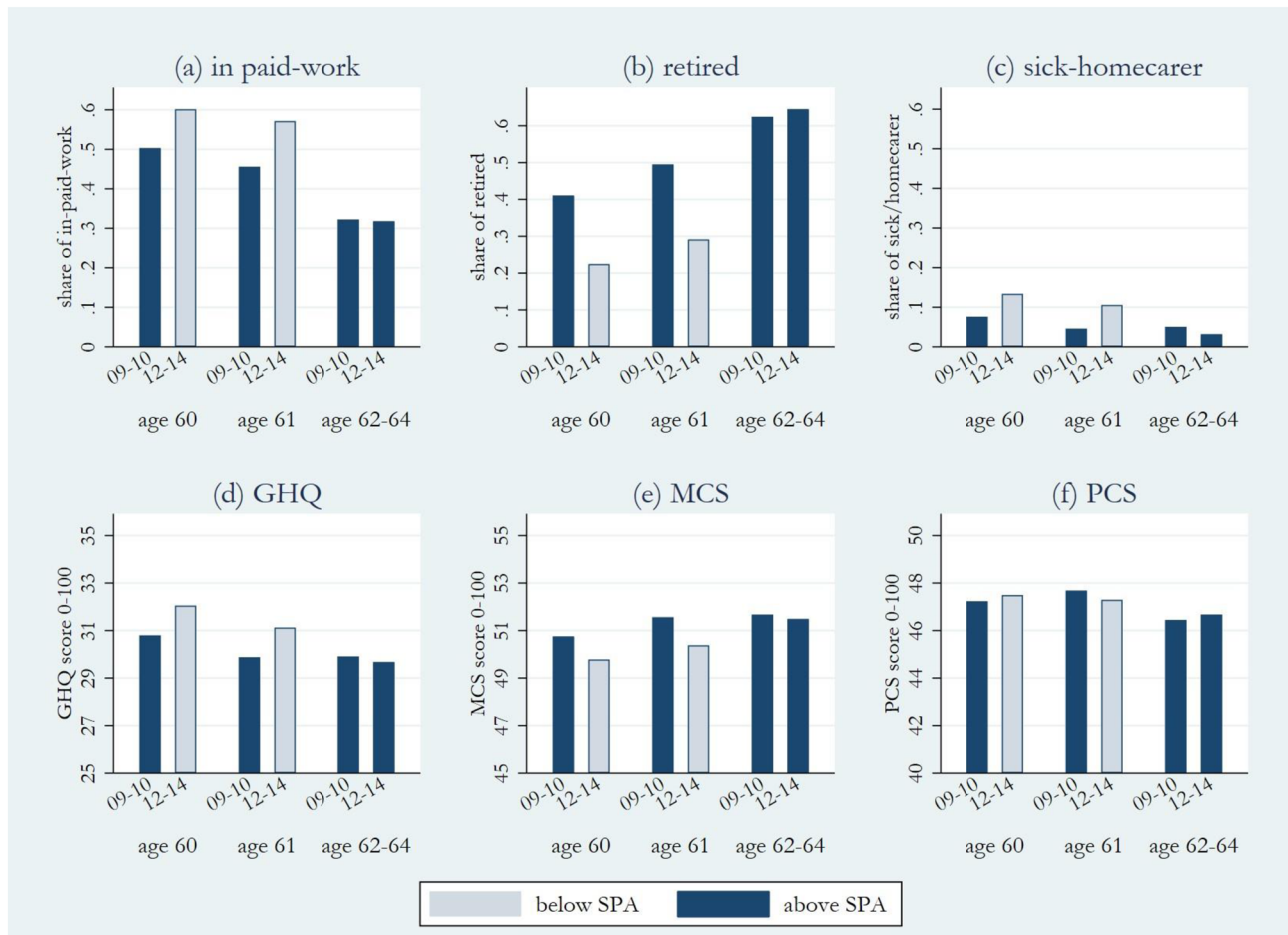
<sup>b</sup>*N* = 11,488.

\* < 0.10.

\*\* < 0.05.

\*\*\* < 0.01.





**FIGURE 2** Employment status and health scores in years 2009–2010 and 2012–2014 for women aged 60, 61, and 62–64 years.

*Note.* we show average values for each age-group. All respondents are above-statutory pensionable age (SPA) in 2009–2010. Women aged 60 and 61 years are below-SPA in 2012–2014 due to the reform. Higher General Health Question (GHQ; mental health score (MCS)/physical health scores (PCS)) scores signal worse (better) health [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Column 2 indicates that being below the SPA due to the reform leads to a significant increase of 1.9 points in GHQ depression scores. Evaluated at the sample-average GHQ score of 30.4, this corresponds to an elasticity of 6.5%. The ES amounts to 14% of the GHQ standard deviation. Although similar ES have been defined as sizeable effects in recent studies in economics (Dustmann & Fasani, 2016), they would not be considered meaningful under the MID rationale (see Section 3.1). On the other hand, we estimate that the SPA reform increases the likelihood of suffering from common mental disorders (GHQ cutoff) by 6.2 percentage points (elasticity of 17%, Column 3). Negative effects (not statistically significant at 10%) are found for both the mental- and physical-health SF-12 scores.

Table 3 explores the impact of the length of SPA postponement on health (model (3.3)). Confirming previous results, it suggests that a longer postponement of SPA leads to worse GHQ and MCS scores: for example, relative to women unaffected by the reform, the GHQ-score increases by 1.96 points for those with an SPA increase of 6–24 months (elasticity 6.4%) and by 3.1 for those with an increase of 36 months or more (+10.1%). The latter effect exceeds the MID threshold (small effect, ES = 0.21). A similar result is found for the likelihood of common mental disorders (GHQ-cut-off), and for the MCS, although the latter would not constitute a MID. No clear pattern emerges for PCS scores.

#### 4.1.1 | SES heterogeneity

In Table 4 we show the net impact of being below-SPA for routine (Column ii), intermediate (iii) and managerial (iv) workers, based either on their current or last occupation (model (3.2)), alongside the results from the baseline model (3.1) in Column (i), and the sample average for each outcome. The effect of the reform on employment did not

**TABLE 2** Effect of State Pension age postponement on employment, GHQ, MCS, and PCS index

	(1) In paid work	(2) GHQ	(3) GHQ cutoff	(4) MCS	(5) PCS
Being below-SPA	0.101*** (0.035)	1.993** (1.032)	0.062** (0.032)	−0.773 (0.614)	−0.564 (0.857)
Routine SES (ref.)					
Intermediate SES	−0.005 (0.023)	−0.751 (0.634)	−0.023 (0.021)	0.615* (0.371)	1.322** (0.619)
Manager SES	−0.071*** (0.024)	−1.860*** (0.595)	−0.071*** (0.020)	1.039*** (0.371)	2.692*** (0.651)
Marital status (ref. “married/couple”)					
Widowed/divorced	0.008 (0.019)	3.956*** (0.666)	0.119*** (0.019)	−3.237*** (0.408)	−3.269*** (0.535)
Single	−0.090*** (0.034)	0.296 (1.124)	0.010 (0.037)	−1.710** (0.819)	−1.833* (1.079)
No children (ref.)					
1–2 children	−0.078*** (0.021)	−0.905 (0.625)	−0.032 (0.020)	0.224 (0.392)	−0.728 (0.484)
3+ children	−0.073*** (0.027)	0.736 (0.736)	0.010 (0.025)	−0.623 (0.441)	−2.250*** (0.656)
No education (ref.)					
Low education	0.127*** (0.023)	−0.816 (0.722)	−0.031 (0.023)	1.028** (0.460)	2.966*** (0.628)
Mid or high education	0.194*** (0.025)	−2.134*** (0.809)	−0.053** (0.023)	1.857*** (0.526)	3.676*** (0.676)
Constant	2.087*** (0.767)	11.284 (29.883)	−0.371 (0.899)	74.904*** (18.157)	34.826 (24.049)
N observations	7,374	7,374	7,374	7,374	7,374
N women	3,531	3,531	3,531	3,531	3,531
Mean	0.435	30.4	0.33	51	46.8
Standard deviation	0.49	14	0.47	9.4	12.2

Note. Columns 1 and 3 report Linear Probability Model estimates for being in paid-work (yes/no) and being above the GHQ cutoff of 12+ (0–36 scale; Goldberg et al., 1997); Columns 2, 4, and 5 report OLS estimates for the GHQ, MCS, and PCS index (0 – *least distressed*; 100 – *most distressed*). All estimates refer to Model (3.1). The status of being above/below SPA is defined by comparing the individual SPA (based on month-year of birth) and the date of interview. The sample includes women aged 60–64 between 2009 and 2016, observed either above-SPA (never affected by the reform) or below-SPA (affected by the reform), having been engaged in paid-work in their life. Additional controls include fixed effects for age (in quarters), interview year (in quarters), and country and a linear control for year-and-month-of-birth. Standard errors are clustered by year-and-month-of-birth (152 clusters).

Abbreviations: GHQ, General Health Question; MCS, mental health score; PCS, physical health score; SES, socioeconomic status; SPA, State Pension age.

\* < 0.10.

\*\* < 0.05.

\*\*\* < 0.01.

differ systematically by SES: among routine SES women, the reform increased employment by 10 percentage points. This effect is not statistically different than the employment effect for intermediate (+13.5) and managerial (+8) SES groups.

The negative effects of the SPA reform on aggregate measures of health is significantly stronger for women in the lowest occupations and not statistically significant for higher SES workers in managerial occupations. For routine workers, being below-SPA significantly increases GHQ depression scores and reduces physical health SF-12 scores. The effects observed for mental health are of clinical relevance: we estimate an increase of 3.5 points in the GHQ score (elasticity 11.4%), which exceeds the MID cutoff (small effect, ES = 0.25), and a 10.4-percentage point increase in the probability of depression based on the GHQ cut-off. Moreover, results from the analysis on the specific GHQ and SF-12 factors suggest that the lowest SES experience a significant decline in all dimensions of mental health. We also observe a statistically significant 2.4-point reduction in the PCS score (5.5%); although this effect on physical health is clinically relevant under the ES definition (small effect, ES = 0.2), it is not minimally important according to the anchoring approach (cutoff of 6.8 points).

#### 4.1.2 | The wear-and-tear effect

We now investigate whether heterogeneity in the health effect of the reform can be explained by prolonged exposure to jobs characterised by different levels of demand and control at work. This is often referred to as the “wear-and-tear” effect of work, whereby each occupation carries a different level of physical and psychosocial occupational stress (Karasek, 1979; Ravesteijn et al., 2018). We employ a job-exposure matrices (JEM) recently built by Kroll and Lampert (2011) based on a large representative survey on working conditions in Germany. The JEM provides a

**TABLE 3** Effect of State Pension age extent of postponement on GHQ, MCS, and PCS indices

	(6) GHQ	(7) GHQ cutoff	(8) MCS	(9) PCS
0 postponement (ref.)				
1–6 months	0.051 (1.706)	−0.003 (0.060)	−0.114 (1.013)	0.363 (1.597)
6–24 months	1.963* (1.072)	0.069* (0.036)	−0.579 (0.684)	−0.569 (1.033)
24–36 months	2.318* (1.311)	0.061* (0.037)	−1.126 (0.773)	−0.635 (1.138)
36+ months	3.068** (1.508)	0.087** (0.044)	−1.423* (0.844)	−0.428 (1.509)
Routine job (reference)				
Intermediate	−0.768 (0.636)	−0.024 (0.021)	0.620* (0.371)	1.320** (0.618)
Manager	−1.867*** (0.596)	−0.071*** (0.020)	1.041*** (0.371)	2.690*** (0.651)
Marital status (ref. “married/couple”)				
Widowed/divorced	3.963*** (0.666)	0.120*** (0.019)	−3.240*** (0.408)	−3.269*** (0.535)
Single	0.287 (1.127)	0.010 (0.037)	−1.704** (0.822)	−1.828* (1.079)
No children (ref.)				
1–2 children	−0.907 (0.625)	−0.032 (0.020)	0.223 (0.391)	−0.728 (0.484)
3+ children	0.744 (0.736)	0.010 (0.025)	−0.626 (0.441)	−2.251*** (0.656)
No education (ref.)				
Low education	−0.812 (0.721)	−0.031 (0.023)	1.024** (0.459)	2.963*** (0.629)
Mid or high education	−2.125*** (0.811)	−0.053** (0.023)	1.848*** (0.527)	3.676*** (0.677)
Constant	9.699 (30.030)	−0.408 (0.912)	75.915*** (18.207)	34.393 (24.240)
N observations	7,374	7,374	7,374	7,374
N women	3,531	3,531	3,531	3,531
Mean	30.4	0.33	51	46.8
Standard deviation	14	0.47	9.4	12.2

Note. Columns 6, 8, and 9 report OLS estimates for the GHQ, MCS and PCS index (0 - *least distressed*; 100 - *most distressed*); Column 2 reports Linear Probability Model estimates for being above the GHQ cutoff of 12+ (0–36 scale; Goldberg et al., 1997). All estimates refer to Model (3.3). The SPA postponement is measured as a difference in months between the individual-specific SPA postreform and the prereform threshold of 60 years old. The sample includes women aged 60–64 between 2009 and 2016, observed either above-SPA (never affected by the reform) or below-SPA (affected by the reform), having been engaged in paid-work in their life. Additional controls include fixed effects for age (in quarters), interview year (in quarters), and country and a linear control for year-and-month-of-birth. Standard errors are clustered by year-and-month-of-birth (152 clusters).

Abbreviations: GHQ, General Health Question; MCS, mental health score; PCS, physical health score; SPA, State Pension age.

\* < 0.10,

\*\* < 0.05.

\*\*\* < 0.01.

dichotomous index of job demand, which summarises five dimensions of occupational burden: Ergonomic Stress, Environmental Pollution, Mental Stress, Social Stress, and Temporal Loads. This measure has been externally validated (see Santi, Kroll, Dietz, Becher, and Ramroth (2013)) and recently applied in economics research (Mazzonna & Peracchi, 2017). We complement this measure with dichotomous index of high/low job control built by Solovieva et al. (2014) from a large survey on adult Finnish workers. The index summarises the degree of decision authority and skill discretion and has been externally validated. Details are included in Appendix 1.3. We were able to match both indices to 99% of our sample through respondents' current or last 4-digit ISCO code. Around 33% of the sample has (had) a high-strain job, that is, highly physically or psychosocially demanding, with low authority and discretion. This includes, for example, housekeeping and restaurant services, personal-carers, salespersons, cleaners, and machine operators. Among them, 85% belong to the routine SES (see Appendix 1.3).

We estimate model (3.1) interacting the SPA-eligibility, job-demand, and job-control indicators.<sup>7</sup> Results (Table 5) provide evidence that a postponement of SPA has a significant negative impact only on the health of women in high straining occupations (high-demand, low-control): GHQ depression score increases by 3.8 points (27% of the variable's standard deviation) with a mean elasticity of 13%, MCS score drops by 1.1 points (12% of standard deviation, elasticity of 2.1%, only significant at 13%), and PCS drops by 2.1 points (17% of standard deviation, elasticity of 4.5%). The result

<sup>7</sup>We find no evidence of selection into nonstraining jobs due to the statutory pensionable age (SPA) change. Running models like (3.1) for the probability of having a straining job, no effect of pension eligibility emerges (available upon request).

**TABLE 4** Heterogeneous effect of State Pension age postponement by SES

	Mean (SD)	Effect of being below-SPA			
		(i) Overall effect	(ii) Routine SES	(iii) Intermediate SES	(iv) Managerial SES
Share of $N = 7,374$			38.6%	28.1%	33.3%
In paid work	0.43 (0.49)	0.101 <sup>***</sup> (0.035)	0.098 <sup>**</sup> (0.04)	0.135 <sup>***</sup> (0.045)	0.08 <sup>*</sup> (0.043)
Health indices					
GHQ score	30.4 (14)	1.993 <sup>**</sup> (1.032)	3.468 <sup>**</sup> (1.367)	1.039 (1.201)	1.179 (1.062)
GHQ cutoff	0.33	0.062 <sup>**</sup> (0.032)	0.104 <sup>**</sup> (0.040)	0.046 (0.037)	0.029 (0.035)
MCS score	51.0 (9.4)	−0.773 (0.614)	−1.059 (0.805)	−0.777 (0.743)	−0.478 (0.663)
PCS score	46.9 (12.2)	−0.564 (0.857)	−2.446 <sup>**</sup> (1.074)	0.506 (1.142)	0.579 (1.000)
GHQ factors					
Anxiety	27.6 (20.8)	2.491 <sup>*</sup> (1.527)	3.973 <sup>**</sup> (1.903)	0.887 (1.794)	2.138 (1.594)
Social dysfunction	36.6 (11.1)	1.400 <sup>*</sup> (0.849)	2.588 <sup>**</sup> (1.082)	0.855 (1.008)	0.586 (0.887)
Confidence loss	17.5 (21.2)	2.773 <sup>**</sup> (1.380)	5.099 <sup>**</sup> (1.969)	1.897 (1.604)	1.04 (1.512)
SF-12 factors					
Physical function	48.0 (11.2)	−0.573 (0.748)	−1.926 <sup>**</sup> (0.963)	0.081 (0.997)	0.332 (0.845)
Role physical	48.6 (10.7)	−0.600 (0.729)	−2.043 <sup>**</sup> (0.928)	0.027 (0.943)	0.416 (0.811)
Bodily pain	47.4 (12.3)	−0.561 (0.814)	−2.489 <sup>**</sup> (1.023)	0.192 (1.074)	0.856 (0.961)
General health	46.1 (12.6)	−0.935 (0.836)	−2.960 <sup>***</sup> (1.042)	0.887 (1.122)	−0.187 (0.985)
Vitality	50.2 (10)	−0.668 (0.634)	−0.954 (0.797)	−0.522 (0.812)	−0.482 (0.712)
Social functioning	50.1 (10.4)	−0.697 (0.597)	−1.494 <sup>*</sup> (0.851)	−0.665 (0.804)	0.09 (0.695)
Role emotional	49.8 (11.3)	−0.775 (0.619)	−1.846 <sup>**</sup> (0.847)	−0.278 (0.811)	−0.044 (0.643)
Mental	50.1 (15.7)	−0.755 (0.652)	−1.267 (0.843)	−0.618 (0.771)	−0.334 (0.668)

*Note.* We show coefficients for “being below-SPA” from Model (3.1) in Column i, and from Model (3.2) in Columns ii–iv, where we report the net effect of being below-SPA for routine, intermediate, and managerial workers. Controls include fixed effects for age (in quarters), interview year (in quarters), SES, living arrangements and marital status (married/in couple, widowed/divorced/separated, single), country, number of children (0, 1–2, 3+), education (low, mid, or high degree), and a linear control for year-and-month of birth. The status of being above/below SPA is defined by comparing the individual SPA (based on month-year of birth) and the date of interview. Standard errors are clustered by year-and-month of birth. The sample includes women aged 60–64 between 2009 and 2016, observed either above-SPA (never affected by the reform) or below-SPA (affected by the reform), having engaged in paid-work in their life.

Abbreviations: GHQ, General Health Question; MCS, mental health score; PCS, physical health score; SES, socioeconomic status; SF-12, Short-Form-12; SPA, State Pension age.

\* < 0.10.

\*\* < 0.05.

\*\*\* < 0.01.

for GHQ is a meaningful effect according to the MID criterion, with an ES of 0.27 *SD*. The probability of being clinically depressed (based on the GHQ cut-off) increases by 12 percentage points (average prevalence 33%). Finally, it is worth noting that the increase in labour market attachment is similar in magnitude and not statistically different between the group that is exhibiting a health decline and those who are not. Thus, results do not seem to support the hypothesis that the stronger effect on health for women in routine SES (or in high-strain jobs) is due to a stronger incentive to remain employed.

## 4.2 | Robustness checks

### 4.2.1 | Diagnosed diseases

We investigate whether the observed health decline triggered the onset of chronic conditions, which could significantly impact mortality and health care expenditure (Behncke, 2012). Given the recent implementation of the reform, we focus on conditions which can plausibly be affected in the short term (Table 6). For each disease, we exclude respondents who were diagnosed with the condition before entering the study (Moon, Glymour, Subramanian, Avendaño, and Kawachi (2012)). We estimate model (3.1) with a dichotomous dependent variable for

**TABLE 5** Heterogeneous effect of State Pension age postponement by job demand and job control level

	Mean (SD)	Effect of being below-SPA			
		(ii) High control-low demand	(iii) Low control-high demand	(iv) Low control-low demand	(v) High control-high demand
Share of $N = 7,374$		29%	33%	27%	11%
In paid work	0.43 (0.49)	0.092** (0.044)	0.100*** (0.040)	0.140*** (0.044)	0.099* (0.054)
Health indices					
GHQ score	30.4 (14)	1.342 (1.264)	3.829*** (1.392)	1.183 (1.213)	1.150 (1.612)
GHQ cutoff	0.33	0.047 (0.039)	0.123*** (0.040)	0.035 (0.038)	-0.008 (0.055)
MCS score	51.0 (9.4)	-0.602 (0.736)	-1.132 (0.853)	-0.833 (0.769)	-0.363 (0.956)
PCS score	46.9 (12.2)	-0.023 (1.010)	-2.092** (1.047)	0.211 (1.095)	0.401 (1.488)
GHQ factors					
Anxiety	27.6 (20.8)	1.986 (1.894)	4.331** (1.885)	1.389 (1.791)	1.614 (2.375)
Confidence loss	36.6 (11.1)	0.809 (1.015)	2.689** (1.139)	1.113 (0.962)	1.193 (1.196)
Social dysfunction	17.5 (21.2)	1.815 (1.558)	6.210*** (2.002)	1.212 (1.851)	0.315 (2.546)
SF-12 factors					
Physical functioning	48.0 (11.2)	-1.143 (1.440)	-2.179** (0.982)	0.005 (0.832)	0.843 (1.210)
Role physical	48.6 (10.7)	-1.098 (1.325)	-2.234** (0.970)	0.171 (0.755)	0.843 (1.178)
Bodily pain	47.4 (12.3)	-0.976 (1.574)	-2.879*** (1.060)	0.328 (0.866)	1.225 (1.313)
General health	46.1 (12.6)	-1.236 (1.603)	-3.479*** (1.046)	0.046 (0.933)	0.781 (1.489)
Vitality	50.2 (10)	-0.259 (1.183)	-1.083 (0.810)	-0.548 (0.677)	-0.244 (1.029)
Social functioning	50.1 (10.4)	0.602 (1.135)	-2.107** (0.919)	-0.586 (0.597)	0.653 (1.128)
Role emotional	49.8 (11.3)	0.221 (1.140)	-2.384*** (0.911)	-0.533 (0.618)	0.7 (1.012)
Mental	50.1 (15.7)	-0.159 (1.306)	-1.520* (0.883)	-0.794 (0.695)	0.241 (0.905)

*Note.* We report OLS coefficients from a model based on (3.2), where we added a three-way interaction term between SPA eligibility (below-SPA), high/low job control, and high/low job demand, which are also added as separate covariates. Columns i-iv report the net effect of being below-SPA for routine (and nonroutine) groups with high- or low-demand jobs. Additional controls include socioeconomic status, fixed effects for age (in quarters), interview year (in quarters), living arrangements and marital status (married, widowed/divorced/separated, single), country, number of children (none, one-two, three, or more), education (low, mid, or high degree), and a linear control for year-and-month of birth. The routine classification follows the National Statistics SEC-3 taxonomy. The status of being above/below SPA is defined by comparing the individual SPA (based on month-year of birth) and the date of interview. High job demand was built and validated by Santi et al. (2013); the job control measure was built and validated by Solovieva et al. (2014); both are linked to respondents' ISCO code. Standard errors are clustered by year-and-month-of-birth. The sample includes women aged 60–64 between 2009 and 2016, observed either above-SPA (never affected by the reform) or below-SPA (affected by the reform), having engaged in paid-work in their life.

Abbreviations: GHQ, General Health Question; MCS, mental health score; PCS, physical health score; SPA, State Pension age.

\* $p < 0.10$ .

\*\* $p < 0.05$ .

\*\*\* $p < 0.01$ .

having been newly diagnosed with the disease since the previous interview. We interact the pension-eligibility dummy with the job-control and the job-demand dummies. Results in Table 6 show a statistically significant increase only in the probability of a doctor's diagnosis of clinical depression (+1.5 percentage-points), only among women with high-demand and low-control jobs (average incidence of new depression diagnoses = 1.1%, average baseline prevalence of diagnosed-depression = 5.9%).<sup>8</sup>

#### 4.2.2 | Placebo tests

We run a falsification test on the male population, whose work status should not be directly affected by the reform. Although males might adjust their retirement decision as a result of their wife's retirement age, we would expect to see

<sup>8</sup>Due to data limitations, we cannot evaluate changes in healthy behaviours (e.g., drinking/smoking/exercising).



**TABLE 6** Effect of being below-SPA on new diagnosis of chronic conditions, by SES and job demand level

	Average incidence of new diagnoses	N	Effect of being below-SPA			
			(ii) High control-low demand	(iii) Low control-high demand	(iv) Low control-low demand	(v) High control-high demand
Arthritis	0.048	3,428	0.005 (0.019)	0.032 (0.022)	−0.014 (0.021)	−0.012 (0.026)
Coronary heart disease	0.004	4,852	0.002 (0.003)	0.002 (0.005)	0.001 (0.003)	0.001 (0.006)
Angina	0.004	4,779	0.007 (0.004)	0.006 (0.005)	0.005 (0.004)	0.009 (0.008)
Heart attack	0.002	4,810	−0.001 (0.001)	−0.002 (0.002)	−0.001 (0.001)	0.003 (0.003)
Liver condition	0.004	4,850	0.002 (0.005)	−0.004 (0.004)	0.001 (0.007)	0.002 (0.006)
Diabetes	0.01	4,588	0.004 (0.007)	−0.001 (0.007)	0.018 (0.006)	0.012 (0.010)
High blood pressure	0.043	3,727	−0.004 (0.015)	−0.028 (0.020)	−0.002 (0.019)	−0.002 (0.022)
Clinical depression	0.011	4,643	0.007 (0.006)	0.015** (0.007)	0.006 (0.008)	0.013 (0.012)

*Note.* We report OLS coefficients from Model (3.1) where we added a three-way interaction term between SPA eligibility (below-SPA), high/low job control, and high/low job demand, which are also added as separate covariates. Columns i–iv report the net effect of being below-SPA on incidence of new diagnoses, for routine (and nonroutine) groups with high- or low-demand jobs. Additional controls include fixed effects for age (in quarters), interview year (in quarters), living arrangements and marital status (married, widowed/divorced/separated, single), country, number of children (none, one–two, three or more), education (low, mid, or high degree), and a linear control for year-and-month of birth. The routine classification follows the National Statistics SEC-3 taxonomy. The status of being above/below SPA is defined by comparing the individual SPA (based on month-year of birth) and the date of interview. High job-demand is measured through the JEM by Santi et al. (2013); job-control is measured through the JEM by Solovieva et al. (2014); both are linked to respondents' ISCO code. Standard errors are clustered by year-and-month of birth. The sample includes women aged 60–64 years between 2009 and 2016, observed either above-SPA (never affected by the reform) or below-SPA (affected by the reform), having been engaged in paid-work in their life, excluding those reporting to have been diagnosed with the specific disease at baseline. The share of respondents in the main sample (7,374 obs.) who, at baseline, had already been diagnosed for a specific condition is as follows: arthritis 0.311, coronary heart disease 0.012, angina 0.028, heart attack 0.019, liver condition 0.012, diabetes 0.069, high blood pressure 0.251, and clinical depression 0.059.

Abbreviations: JEM, job-exposure matrices; SES, socioeconomic status; SPA, State Pension age.

\* $p < 0.10$ .

\*\* $p < 0.05$ .

\*\*\* $p < 0.01$ .

weaker effects for men than for women. After assigning women's SPA to men, based on their birthdate, with find no effect of pension eligibility on male's employment or health (Table 7). We then focus on women who never engaged in paid work and who are unlikely to be induced to work by the SPA postponement. Hence, the reform should not affect this population's health. Due to a small sample (1,410 women aged 60–64 years), we have limited statistical power. However, results suggest that being below-SPA does not affect any health index among those women (Table 7).

#### 4.2.3 | Income effect

Previous research has found that the SPA-change reduced after-tax individual income for women by 20%, it had a smaller effect on household income (−6%), and it increased absolute poverty rates between 6 and 8 percentage points, yet it had no impact on material deprivation (Cribb et al., 2016; Cribb & Emmerson, 2018). In our sample, we estimate that the SPA-reform reduced individual and household median after-tax income by 12% and 6.7%, respectively, and it increased absolute poverty rates by 6 percentage points (baseline poverty rate among untreated women is 11.7%). Full details are in Appendix 1.5.

A reduction in income may translate into negative mental health consequences, for example, by reducing the ability to afford basic goods or increasing the likelihood of individual and household indebtedness (Keese and Schmitz, 2014). We thus re-estimate our models by including household or individual log-income as additional regressor to examine whether this might affect the results (Table 8, Panel 1; models controlling for individual income are available on request). The results are very similar to those in models that do not control for income. Although not a conclusive test, this suggests that the negative effects of the reform on mental and physical health might not only be attributable to the reduction in income.



TABLE 7 Effect of State Pension age postponement on men and on women who never worked

Sample	(0) Paid work		(1) GHQ		(2) MCS		(3) PCS		(4) GHQ		(5) MCS		(6) PCS	
	Men		Men		Men		Men		Women never worked		Women never worked		Women never worked	
Being below-SPA	-0.014 (0.041)		-0.501 (0.990)		-0.468 (0.618)		-0.152 (0.904)		-1.506 (1.937)		0.156 (1.286)		-1.631 (1.826)	
Routine job (reference)														
Intermediate	0.069 <sup>**</sup> (0.026)		-0.074 (0.684)		0.795 <sup>*</sup> (0.436)		1.383 <sup>**</sup> (0.536)		-		-		-	
Manager	-0.022 (0.023)		-1.393 <sup>***</sup> (0.528)		1.035 <sup>***</sup> (0.385)		3.597 <sup>***</sup> (0.558)		-		-		-	
Marital status (ref. "married/couple")														
Widowed/divorced	-0.137 <sup>***</sup> (0.029)		3.074 <sup>***</sup> (0.738)		-2.827 <sup>***</sup> (0.472)		-3.026 <sup>***</sup> (0.686)		4.203 <sup>**</sup> (1.769)		-3.864 <sup>***</sup> (1.090)		-4.166 <sup>***</sup> (1.456)	
Single	-0.252 <sup>***</sup> (0.036)		2.208 <sup>**</sup> (0.854)		-1.804 <sup>***</sup> (0.590)		-2.149 <sup>**</sup> (0.842)		1.426 (3.033)		-1.214 (1.462)		-3.647 (2.602)	
No children (ref.)														
1-2 children	-0.103 <sup>***</sup> (0.022)		1.435 <sup>**</sup> (0.556)		-0.402 (0.352)		-1.286 <sup>***</sup> (0.488)		-3.199 <sup>**</sup> (1.491)		2.560 <sup>**</sup> (1.073)		1.076 (1.638)	
3+ children	-0.089 <sup>***</sup> (0.027)		2.826 <sup>***</sup> (0.643)		-1.635 <sup>***</sup> (0.399)		-2.536 <sup>***</sup> (0.579)		5.031 <sup>**</sup> (2.100)		-2.789 <sup>*</sup> (1.623)		-5.080 <sup>***</sup> (1.448)	
No education (ref.)														
Low education	0.065 <sup>*</sup> (0.036)		-2.269 <sup>***</sup> (0.734)		1.696 <sup>***</sup> (0.498)		2.898 <sup>***</sup> (0.857)		-1.650 (1.540)		1.468 <sup>*</sup> (0.849)		4.647 <sup>***</sup> (1.514)	
Mid or high education	0.051 <sup>*</sup> (0.030)		-3.607 <sup>***</sup> (0.667)		1.944 <sup>***</sup> (0.498)		4.822 <sup>***</sup> (0.807)		-6.706 <sup>***</sup> (1.354)		3.909 <sup>***</sup> (0.877)		6.892 <sup>***</sup> (1.343)	
N observations	6,122		6,122		6,122		6,122		1,410		1,410		1,410	
N women	3,125		3,125		3,125		3,125		615		615		615	

Note. We report OLS estimates for the GHQ index (ranging from 0 – least distressed to 100 – most distressed), MCS and PCS index (ranging from 0 – most distressed to 100 – least distressed). Estimations 1–3 are based on a sample of men aged 60–64 years between 2009 and 2016, having been engaged in paid work in their life. Estimations 4–6 are based on a sample of women aged 60–64 years who never engaged in paid work. The status of being above/below SPA is defined by comparing the individual SPA (based on month-year of birth, assuming that women's rules apply to men) and the date of interview. Additional controls include fixed effects for age (in quarters), interview year (in quarters), and country and a linear control for year-and-month of birth. Standard errors are clustered by year-and-month of birth.

Abbreviations: GHQ, General Health Question; MCS, mental health score; PCS, physical health score; SPA, State Pension age.

\* <0.10.

\*\* <0.05.

\*\*\* <0.01.

**TABLE 8** Effect of State Pension age postponement on health, sensitivity analysis

	Effect of being below-SPA				
	(i) Overall effect	(ii) High control-low demand	(iii) Low control-high demand	(iv) Low control-low demand	(v) High control-high demand
(1) Including after-tax household log-income as regressor ( $N = 7,374$ )					
GHQ score	2.101** (1.035)	1.309 (1.247)	3.817*** (1.423)	1.331 (1.221)	1.480 (1.650)
GHQ cut-off	0.055* (0.032)	0.038 (0.038)	0.116*** (0.042)	0.035 (0.039)	-0.004 (0.058)
MCS score	-0.921 (0.629)	-0.680 (0.737)	-1.189 (0.878)	-0.884 (0.791)	-0.540 (0.974)
PCS score	-0.553 (0.877)	0.052 (0.981)	-1.943** (1.060)	0.157 (1.123)	0.442 (1.493)
(2) Standard errors clustered at individual level (3,494 clusters, $N = 7,374$ )					
GHQ score	1.993* (1.054)	1.342 (1.236)	3.829*** (1.314)	1.183 (1.323)	1.150 (1.619)
GHQ cut-off	0.062* (0.033)	0.047 (0.039)	0.123*** (0.040)	0.035 (0.041)	-0.008 (0.053)
MCS score	-0.773 (0.676)	-0.602 (0.771)	-1.132 (0.857)	-0.833 (0.845)	-0.363 (1.071)
PCS score	-0.564 (0.926)	-0.023 (1.052)	-2.092** (1.161)	0.211 (1.152)	0.401 (1.474)
(3) Alternative age/time control: linear specification for age-quarter and year-quarter ( $N = 7,374$ )					
GHQ score	2.170** (1.030)	1.524 (1.247)	4.078*** (1.372)	1.388 (1.251)	1.230 (1.602)
GHQ cut-off	0.055* (0.031)	0.042 (0.038)	0.112*** (0.040)	0.029 (0.039)	-0.018 (0.056)
MCS score	-0.867 (0.597)	-0.680 (0.722)	-1.28* (0.835)	-0.911 (0.766)	-0.341 (0.931)
PCS score	-0.814 (0.857)	-0.333 (0.980)	-2.3** (1.047)	-0.033 (1.092)	0.067 (1.499)
(4) Alternative age/time control: quadratic specification for age-quarter and year-quarter ( $N = 7,374$ )					
GHQ score	1.968* (1.023)	1.310 (1.239)	3.892*** (1.384)	1.196 (1.232)	1.020 (1.573)
GHQ cut-off	0.056* (0.032)	0.043 (0.038)	0.121*** (0.040)	0.03 (0.039)	-0.017 (0.056)
MCS score	-0.774 (0.612)	-0.587 (0.734)	-1.2 (0.848)	-0.828 (0.773)	-0.246 (0.946)
PCS score	-0.620 (0.865)	-0.132 (1.004)	-2.13** (1.051)	0.144 (1.099)	0.283 (1.489)

Columns ii–v report the net effect of being below-SPA for women with high or low job control and with high or low job demand.

Abbreviations: GHQ, General Health Question; MCS, mental health score; PCS, physical health score; SPA, State Pension age.

\* $p < 0.10$ .

\*\* $p < 0.05$ .

\*\*\* $p < 0.01$ .

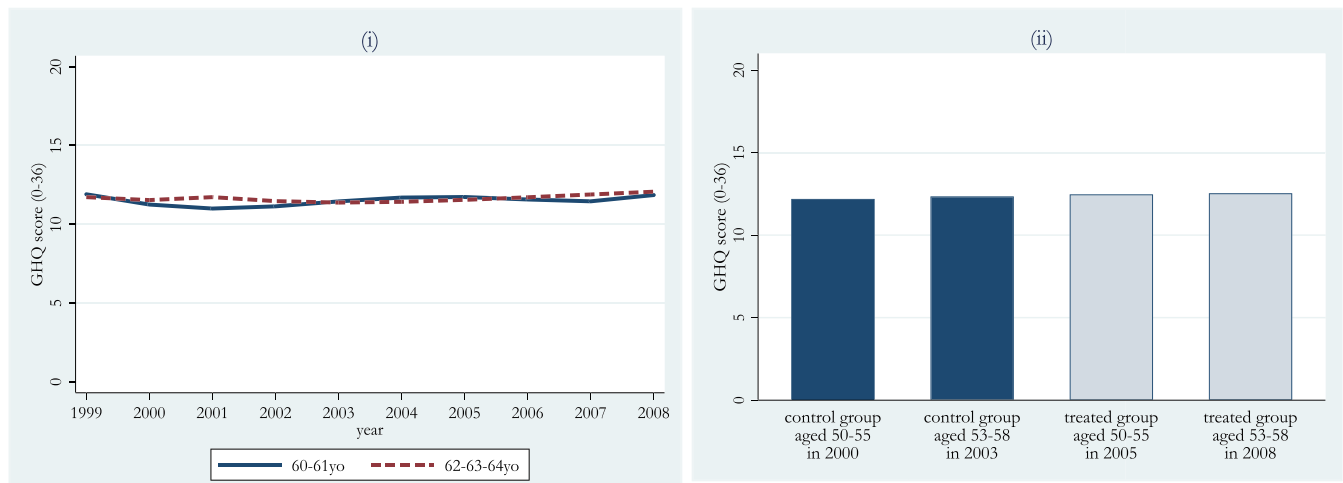
#### 4.2.4 | Econometric specification and common trend assumption

As our sample includes repeated observations for some individuals, we test that our results are robust to clustering the standard errors at the individual level (Table 8, Panel 2). Results are also robust to alternative specifications for age and time, such as adopting (a) a linear specification or (b) a quadratic specification for age-quarter and year-quarter (Table 8, Panels 3 and 4) or dropping the control for year-month-of-birth (upon request).<sup>9</sup> Moreover, implementing alternative GHQ cut-offs (e.g., Aalto, Elovainio, Kivimäki, Uutela, and Pirkola (2012)) leads to very similar results (available upon request). Our results are robust to alternative sample selections, to controlling for long-term conditions diagnosed prior to the policy, and to controlling for partner's age (Appendix 1.5). Furthermore, our effects are similar for respondents that had been diagnosed with a chronic condition and those who were free of a diagnosis prior to baseline (Appendix 1.5).

Our identification approach assumes that treated and control would have had similar health trends in the absence of the reform. We test whether birth cohorts treated by the reform (birth cohort 1950–1955) had different health levels and trends relative to cohorts unaffected by the SPA change (birth cohort 1944–1949) when at the same age. Using British Household Panel Survey, we look at health outcomes in 2000 and 2003 for the control group (aged 50–55 years in 2000 and 53–58 years in 2003) and for the treatment group in 2005 and 2008 (aged 50–55 years in 2005 and 53–58 years in 2008). Figure 3 Panel (ii) shows no difference in health levels or trends prior to the reform (see Appendix 1.4).

We also examine health trends for treated and control prior to the reform. As our survey started in 2009, we explore data from the British Household Panel Survey (the precursor of Understanding Society, comparable in both sampling and variables collected), focusing on the GHQ score. We first test whether, in the years 1999–2008, GHQ scores were different between women aged 60–61 and 62–64 years (5,041 observations). Both groups are above-SPA between 1999

<sup>9</sup>Our results are robust to allowing for heterogeneous age and time trends by job type (available on request).



**FIGURE 3** common trend evidence for General Health Question (GHQ) scores, by age cohorts (i) and birth cohorts (ii), from British Household Panel Survey (BHPS).

*Note.* data from BHPS. Panel (i) plots average GHQ score (0 to 36) from 1999 to 2008 for women aged 60–61 and 62–64 years in each year (5,041 obs.). Panel (ii) plots GHQ scores for the control group (women born from 1944 to 1949) in 2000 (aged 50 to 55 years) and in 2003 (aged 53 to 58 years); and for the treated group (women born from 1950 to 1955), in 2005 (aged 50 to 55) and 2008 (aged 53 to 58; 2,779 obs.). Appendix 1.4 includes full regressions' results [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

and 2008, while 60–61-year-olds are affected by the SPA change after 2010. GHQ levels and trends were virtually identical for treated and control (see Figure 3 Panel (i) and Appendix 1.4 for details). Furthermore, we show that the two groups do not differ in the prevalence of functional limitations, using the English Longitudinal Survey of Ageing (years 2002–2008, see Appendix 1.4).

Our approach also assumes that age effects are cohort-constant. This assumption may be violated if there is compression of morbidity, that is, cohorts born later have better health at any given age due of postponement of disease or disability (Jagger et al., 2016). This pattern, however, is unlikely to explain our results due to several reasons. First, if there was compression of morbidity, our findings would likely be lower-bound estimates because we find that later born cohorts (exposed to the pension reform) have worse health than cohorts born earlier (unexposed to the reform). If anything, compression of morbidity would lead to underestimation of the true causal effect of the reform. Second, our sample comprises a narrow range of age and cohort groups. This makes the assumption of cohort-constant age effects plausible, as compression of morbidity typically occurs over long time frames. Appendix Figure 5 illustrates the maximum range of birth-years that we observe for each age-group in the sample. As an example, among women aged 60 years, we observe women born between 1948 and 1956. Among women aged 63 years, the birth cohorts range between 1945 and 1950. Hence, the largest cohort differences are at most 8 or 5 years. It is unlikely that such differences would substantially impact the health conditions of women at a given age. We also note that there appears to be no consensus on whether the United Kingdom and other high-income countries are experiencing compression of morbidity. The evidence is, at best, mixed and strongly dependent on method and outcome (Crimmins & Beltrán-Sánchez, 2011; Gondek, Bann, Ning, Grundy, & Ploubidis, 2019; Rechel et al., 2013). For example, a recent study for England covering a 20-years time interval reports compression of morbidity for cognitive impairment and self-perceived health, but dynamic equilibrium for disability (i.e., less severe disability is increasing but more severe disability is not; Jagger et al., 2016).

Although not a conclusive test of morbidity compression, we followed Cribb et al. (2016) and performed an additional robustness test by allowing the age effects to vary by cohort. This is obtained by adding an interaction term between age dummies (in quarters) and birth-year dummies to a set of controls which included dummies for age (in quarters), dummies for time (in quarters), and dummies for birth-year. Because cohort effects are included using birth-year while age and time are measured in quarters, there is no perfect collinearity between the three set of variables. When adding the interaction term between age-quarter and birth-year, identification relies on the fact that women born in the same year may have different pension eligibility depending on their age-quarter at the time of interview: a person aged 62q1 born in 1952q1 would be eligible for pension, while a person aged 62q1 born in 1952q3 would not be eligible. Our results are robust and confirmed under this specification (results available upon request).

## 4.2.5 | Persistency of the effect for affected women

Our models might capture a very short-term increase in depression and anxiety for being “forced” to work longer in high-strain jobs to qualify for the pension (with a lower lifetime income). Data availability prevents us from establishing whether these effects are long-lasting: the first affected cohorts have only been exposed for a few months/years, while the most affected cohorts reach their SPA in 2017 or later. However, we show that the observed health effects do not dissipate over time for women who are still below SPA. We build a categorical variable capturing the amount of time a person has spent below SPA, at the time of interview, with four levels: having been “below SPA” for 1, 2, 3, or more years at the time of the interview (reference is “being above SPA”). We then estimate model (3.1), interacting the exposure variables with a simplified SES measure (low-SES vs. mid-high SES) as follows:

$$y_{iat} = \alpha + \beta_1 \cdot \text{belowSPA}(\text{for 1 year})_{iat} + \beta_2 \cdot \text{belowSPA}(\text{for 2 years})_{iat} + \beta_3 \cdot \text{belowSPA}(\text{for 3 + years})_{iat} + \beta_4 \cdot \text{belowSPA}(\text{for 1 year})_{iat} * \text{routine} + \beta_5 \cdot \text{belowSPA}(\text{for 2 years})_{iat} * \text{routine} + \beta_6 \cdot \text{belowSPA}(\text{for 3 + years})_{iat} * \text{routine} + \beta_7 \cdot \text{routine} + \gamma_a + \eta_t + \delta_c + X'_{iat}\varphi + \varepsilon_{iat}.$$

**TABLE 9** Persistency of health effects for women below SPA

	(6) GHQ	(7) GHQ cut-off	(8) MCS	(9) PCS	N obs
Low SES					2,844
Not affected by reform (ref.)					1,968
$t < 1$ year spent below SPA	2.868** (1.443)	0.109** (0.046)	−0.919 (0.855)	−1.658 (1.041)	461
$1 < t < 2$ years spent below SPA	3.983** (1.535)	0.089** (0.043)	−0.824 (0.907)	−3.549** (1.298)	307
$t > 2$ years spent below SPA	4.426** (2.253)	0.135* (0.073)	−2.502* (1.33)	−3.372** (1.628)	108
Mid-high SES					4,530
Not affected by reform (ref.)	-	-	-	-	2,774
$t < 1$ year spent below SPA	1.001 (1.147)	0.049 (0.039)	−0.769 (0.678)	0.593 (0.847)	955
$1 < t < 2$ years spent below SPA	1.230 (1.114)	0.038 (0.035)	−0.715 (0.705)	0.209 (1.060)	554
$t > 2$ years spent below SPA	0.753 (1.406)	−0.011 (0.047)	0.185 (0.826)	1.385 (1.303)	247

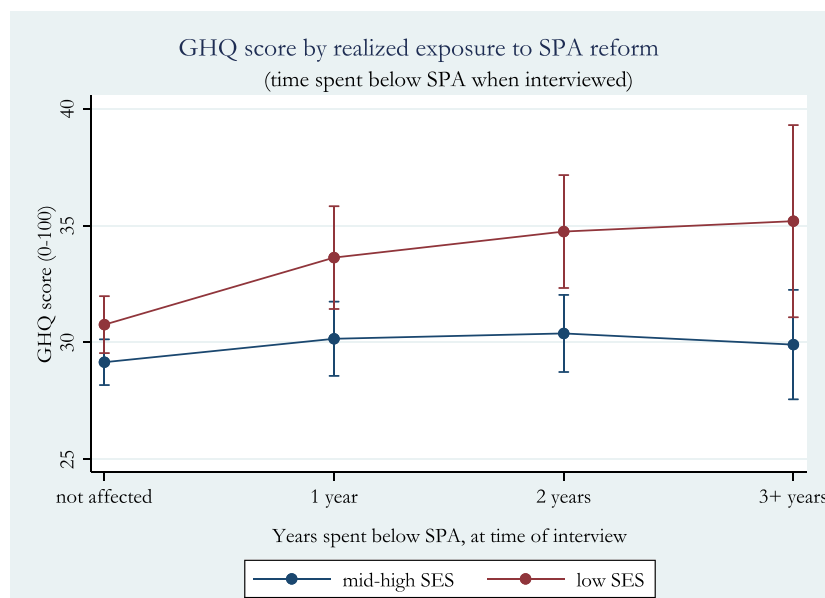
Abbreviations: GHQ, General Health Question; MCS, mental health score; PCS, physical health score; SES, socioeconomic status; SPA, State Pension age.

\* $p < 0.10$ .

\*\* $p < 0.05$ .

\*\*\* $p < 0.01$ .

**FIGURE 4** Persistency of the reform-effect on depression for women below statutory pensionable age (SPA) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



Results in Table 9 (and Figure 4) suggest that the health effects are persistent for women in low SES, and, therefore, do not seem to reflect only a “surprise” effect.

## 5 | DISCUSSION AND CONCLUSIONS

Our results show that the increase in the female State Pension age in the United Kingdom widened health disparities between 2009 and 2016. Specifically, women in lower socioeconomic groups affected by the reform suffered from declines in mental health, which are economically and clinically relevant (a MID), and increased the prevalence of self-reported clinical depression diagnosis. We also found a statistically significant negative effect on physical health for women in the lower socioeconomic group. However, this effect is likely of less clinical relevance.

Our findings are consistent with evidence on—more limited—SPA reforms in the Netherlands (De Grip et al., 2012) and Israel (Shai, 2018) and further suggest that the main mechanism for the observed health deterioration is longer exposure jobs with high demand and low control at work. This can be interpreted in light of theoretical and empirical literature linking straining occupations with health capital degradation, lower productivity, and worse health (Barnay, 2016; Bildt & Michélsen, 2002; Chandola & Zhang, 2017; Fischer & Sousa-Poza, 2009; Paccagnella, 2016; Ravesteijn et al., 2018).

A possible alternative explanation may lie on the detrimental impact of the reform on after-tax individual and household income and on absolute poverty rates (see also Cribb and Emmerson (2018)). This may be partly due to the fact that rates of reemployment are relatively low for older workers due to their higher skill specificity and the risk of age discrimination (OECD, 2017b), especially for routine and manual workers, who are most affected by computerisation and offshoring (Autor & Dorn, 2009). Moreover, the reduction in income is stronger for women in lower socioeconomic groups, which are also those whose health has been most affected by the SPA reform. However, our results are not affected by the inclusion of income measures as covariates, suggesting that the health deterioration might not only be attributable to changes in income.

Our findings may also be attributable to the lack of awareness of women about the reform, as a late and unexpected SPA-change may disrupt pensions-plans, thus generating anxiety and depression (De Grip et al., 2012; Falba, Sindelar, & Gallo, 2009; van Solinge & Henkens, 2017). However, although before 2010 women in lower SES were indeed less aware of the SPA increase than higher SES women (Clery, Humphrey, & Bourne, 2009; MacLeod et al., 2012), recent evidence from Holman, Foster, and Hess (2018) shows that the SPA-change was “almost common knowledge” across all socioeconomic groups by the time the reform came into place (2010/2011), partly due to an extensive public information campaign. Our own findings (Table 3) suggest that the reform's health effects are larger for women facing larger postponements, who turned 60 years old since 2013, when SPA knowledge was widespread. All in all, this evidence does not seem to support lack of awareness as the main explanation for the increase in mental health problems.

Our results estimate the impact of an SPA-increase on health in the months or first years following the introduction of the reform. We are not able to establish whether these effects are long-lasting, as the first cohorts affected by the reform have only been exposed for a short time. Likewise, our models compare outcomes between cohorts exposed to the old-SPA and those that have been for the first time exposed to the new-SPA regime. One might argue that subsequent cohorts might react differently to an increase in the SPA, an assumption that can only be tested in future waves of *Understanding Society*.

Our results are robust to several sensitivity tests, including a falsification test for women who never worked and for men, as well as tests for the common-trends assumption. We also show that violation of the assumption of cohort-constant age effects is unlikely in our sample and would not provide a credible explanation for our results.

Although the State Pension age will continue to increase in OECD countries (OECD, 2017a), our findings suggest that negative health consequences of these reforms have been overlooked and should be considered in cost-effectiveness policy evaluations, as they might outweigh some of the potential benefits from later retirement. Mental illness impacts directly on health-care costs, disability benefit payments and service use, as well as indirectly, for example, by imposing a burden on caregivers and their productivity. Mental disorders hamper labour market productivity (Bubonya, Cobb-Clark, & Wooden, 2017) and are a leading cause of disability worldwide, being associated with the onset of other physical health conditions such as cancer, CVD, and diabetes, which may exacerbate their impact on productivity and the economy (Prince et al., 2007). As a result, mental illness constitutes a major burden for public budgets, estimated at

more than 4% of GDP in the United Kingdom and in OECD countries (Arends, Baer, Miranda, Prinz, & Singh, 2014; OECD/EU, 2018; WHO, 2013).

There are two possible policy implications from our findings. First, the fact that the negative health effects of the reform are confined to women from highly demanding occupations raises potential questions about fairness, and whether eligibility rules should consider occupation as a potential criterion for SPA (Wester & Wolff, 2010). Second, national policies that increase the State Pension age may need to consider strategies to prevent negative health consequences for women in manual and routine occupations, for example, through inclusive labour market policies that facilitate a smooth transition to retirement (OECD, 2017b).

## CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

## ACKNOWLEDGEMENTS

Ludovico Carrino and Mauricio Avendano are supported by the Economic and Social Research Council (ESRC) and the Joint Programme Initiative 'More Years, Better Lives' (ES/P000487/1, Worklong project). Karen Glaser is supported by the Economic and Social Research Council (ESRC, ES/LS002825/1, Wherl Project). This paper represents independent research partly supported by the ESRC Centre for Society and Mental Health at King's College London (ESRC Reference: ES/S012567/1). The views expressed are those of the author(s) and not necessarily those of the ESRC or King's College London. This work was partly supported by European Union Horizon 2020 award "MINDMAP: Promoting Mental Wellbeing in the Ageing Urban Population" (grant number 667661). The study does not necessarily reflect the European Commission's or ESRC's views and in no way anticipates the Commission's or ESRC's future policy in this area. The funding source did not have a role in the design and conduct of the study, the collection, management, analysis, and interpretation of the data, or the preparation, review, approval, or decision to submit the manuscript. The paper makes use of data from Understanding Society, an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by the National Centre for Social Research and TNS BMRB. The research data are distributed by the UK Data Service. University of Essex, Institute for Social and Economic Research, NatCen Social Research and Kantar Public [producers]: Understanding Society: Waves 1-7, 2009-2016 [computer file]. 8th Edition. Colchester, Essex: UK Data Service [distributor], November 2017. SN: 6614 (<https://doi.org/10.5255/UKDA-SN-6614-8>). The authors thank, for comments and suggestion, the Worklong consortium (in particular Lex Burdorf and Merel Schuring), Federico Belotti, Agar Brugiavini, Giorgio Di Gessa, David Lain, Fabrizio Mazzonna, Vahé Nafilyan, Giacomo Pasini, Debora Price, and audiences at the 2017 Boston IHEA conference, the Conference of the Society for Epidemiological Research, the WPEG Conference, the 2018 IAAE Conference in Montreal, the Italian Health Economics Conference, the DIW in Berlin, the London School of Economics, the UCL Institute of Education, the Venice International University, the British Society of Gerontology Conference, the 2019 Geary Institute of Public Policy Seminar series at UCD, the American European Health Economics Workshop in Wien, and the University of Trieste. We also thank the Department for Work and Pensions, Her Majesty's Treasury, AgeUK, and Centre for Ageing Better for useful comments.

## ORCID

Ludovico Carrino  <https://orcid.org/0000-0002-5082-5508>

Karen Glaser  <https://orcid.org/0000-0002-4436-1305>

Mauricio Avendano  <https://orcid.org/0000-0002-7295-2911>

## REFERENCES

- Aalto, A.-M., Elovainio, M., Kivimäki, M., Uutela, A., & Pirkola, S. (2012). The Beck Depression Inventory and General Health Questionnaire as measures of depression in the general population: A validation study using the Composite International Diagnostic Interview as the gold standard. *Psychiatry Research*, 197(1–2), 163–171.
- Arends, I., Baer, N., Miranda, V., Prinz, C., & Singh, S. (2014). *Mental health and work*. Paris: OECD Publishing.
- Atalay, K., & Barrett, G. F. (2014). The causal effect of retirement on health: New evidence from Australian pension reform. *Economics Letters*, 125(3), 392–395.
- Autor, D., & Dorn, D. (2009). This job is "getting old": Measuring changes in job opportunities using occupational age structure. *The American Economic Review*, 99(2), 45–51.



- Barnay, T. (2016). Health, work and working conditions: A review of the European economic literature. *The European Journal of Health Economics*, 17(6), 693–709.
- Behncke, S. (2012). Does retirement trigger ill health? *Health Economics*, 21(3), 282–300.
- Belloni, M., Meschi, E., & Pasini, G. (2016). The effect on mental health of retiring during the economic crisis. *Health Economics*, 25(S2), 126–140.
- Bertoni, M., Brunello, G., & Mazzarella, G. (2018). Does postponing minimum retirement age improve healthy behaviors before retirement? Evidence from middle-aged Italian workers. *Journal of Health Economics*, 58, 215–227.
- Bertoni, M., Maggi, S., & Weber, G. (2018). Work, retirement, and muscle strength loss in old age. *Health Economics*, 27(1), 115–128.
- Bildt, C., & Michélsen, H. (2002). Gender differences in the effects from working conditions on mental health: A 4-year follow-up. *International Archives of Occupational and Environmental Health*, 75(4), 252–258.
- Bloemen, H., Hochguertel, S., & Zweerink, J. (2017). The causal effect of retirement on mortality: Evidence from targeted incentives to retire early. *Health Economics*, 26(12), e204–e218.
- Bonsang, E., Adam, S., & Perelman, S. (2012). Does retirement affect cognitive functioning? *Journal of Health Economics*, 31(3), 490–501.
- Bubonya, M., Cobb-Clark, D. A., & Wooden, M. (2017). Mental health and productivity at work: Does what you do matter? *Labour Economics*, 46, 150–165.
- Bünnings, C., Kleibrink, J., & Weßling, J. (2017). Fear of unemployment and its effect on the mental health of spouses. *Health Economics*, 26(1), 104–117.
- Chandola, T., & Zhang, N. (2018). Re-employment, job quality, health and allostatic load biomarkers: Prospective evidence from the UK Household Longitudinal Study. *International Journal of Epidemiology*, 47(1), 47–57.
- Clark, A. E. (2003). Unemployment as a social norm: Psychological evidence from panel data. *Journal of Labor Economics*, 21(2), 323–351.
- Clery, E., Humphrey, A., & Bourne, T. (2009). Attitudes to pensions: The 2009 survey. *Department of Work and Pension: Research Report*, 701.
- Coe, N., Lindeboom, M. (2008). Does Retirement Kill You? Evidence from Early Retirement Windows. CentER Discussion paper 2008-93.
- Coe, N. B., & Zamorro, G. (2011). Retirement effects on health in Europe. *Journal of health economics*, 30(1), 77–86.
- Coe, N. B., von Gaudecker, H.-M., Lindeboom, M., & Maurer, J. (2012). The effect of retirement on cognitive functioning. *Health Economics*, 21(8), 913–927.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. New York: Routledge.
- Cribb, J., & Emmerson, C. (2017). Can't wait to get my pension: the effect of raising the female state pension age on income, poverty and deprivation. *IFS Working Paper W17/10*.
- Cribb, J., & Emmerson, C. (2018). Can't wait to get my pension: the effect of raising the female early retirement age on income, poverty and deprivation. *Journal of Pension Economics & Finance*, 18(3), 450–472.
- Cribb, J., Emmerson, C., & Tetlow, G. (2016). Signals matter? Large retirement responses to limited financial incentives. *Labour Economics*, 42, 203–212.
- Cribb, J., Hood, A., Joyce, R., & Norris Keiller, A. (2017). *Living standards, poverty and inequality in the UK: 2017*: IFS Report.
- Crimmins, E. M., & Beltrán-Sánchez, H. (2011). Mortality and morbidity trends: Is there compression of morbidity? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66(1), 75–86.
- De Grip, A., Lindeboom, M., & Montizaan, R. (2012). Shattered dreams: The effects of changing the pension system late in the game\*. *The Economic Journal*, 122(559), 1–25.
- Dustmann, C., & Fasani, F. (2016). The effect of local area crime on mental health. *The Economic Journal*, 126(593), 978–1017.
- Eibich, P. (2015). Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of Health Economics*, 43, 1–12. <https://doi.org/10.1016/j.jhealeco.2015.05.001>
- Falba, T., Sindelar, J., & Gallo, W. (2009). Work expectations, realizations, and depression in older workers. *The Journal of Mental Health Policy and Economics*, 12(4), 175–186.
- Farivar, S. S., Liu, H., & Hays, R. D. (2004). Half standard deviation estimate of the minimally important difference in HRQOL scores? *Expert Review of Pharmacoeconomics & Outcomes Research*, 4(5), 515–523.
- Fayers, P. M., & Hays, R. D. (2014). Don't middle your MID: Regression to the mean shrinks estimates of minimally important differences. *Quality of Life Research*, 23(1), 1–4.
- Fischer, J. A. V., & Sousa-Poza, A. (2009). Does job satisfaction improve the health of workers? New evidence using panel data and objective measures of health. *Health Economics*, 18(1), 71–89.
- Galama, T., & Kapteyn, A. (2011). Grossman's missing health threshold. *Journal of Health Economics*, 30(5), 1044–1056.
- Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, 27(1), 191–197.
- Goldberg, D. P., & Williams, P. (1988). *A user's guide to the General Health Questionnaire*. Windsor: NFER-Nelson.
- Gondek, D., Bann, D., Ning, K., Grundy, E., & Ploubidis, G. B. (2019). Post-war (1946–2017) population health change in the United Kingdom: A systematic review. *PLoS ONE*, 14(7), e0218991.
- Graetz, B. (1991). Multidimensional properties of the general health questionnaire. *Social Psychiatry and Psychiatric Epidemiology*, 26(3), 132–138.
- Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2), 223–255.
- Heller-Sahlgren, G. (2017). Retirement blues. *Journal of Health Economics*, 54, 66–78. <https://doi.org/10.1016/j.jhealeco.2017.03.007>

- Hernaes, E., Markussen, S., Piggott, J., & Vestad, O. L. (2013). Does retirement age impact mortality? *Journal of Health Economics*, 32(3), 586–598.
- Holman, D., Foster, L., & Hess, M. (2018). Inequalities in women's awareness of changes to State Pension age in England and the role of cognitive ability. *Ageing and Society*, 144-16140(1), 144–161.
- Jagger, C., Matthews, F. E., Wohland, P., Fouweather, T., Stephan, B. C. M., Robinson, L., ... Brayne, C. (2016). A comparison of health expectancies over two decades in England: Results of the Cognitive Function and Ageing Study I and II. *The Lancet*, 387(10020), 779–786.
- Jayadevappa, R., Cook, R., & Chhatre, S. (2017). Minimal important difference to infer changes in health-related quality of life—A systematic review. *Journal of Clinical Epidemiology*, 89, 188–198. <https://doi.org/10.1016/j.jclinepi.2017.06.009>
- Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, 24(2), 285–308.
- King, M. T. (2011). A point of minimal important difference (MID): A critique of terminology and methods. *Expert Review of Pharmacoeconomics & Outcomes Research*, 11(2), 171–184.
- Kneale, D., & Joshi, H. (2008). Postponement and childlessness: Evidence from two British cohorts. *Demographic Research*, 19, 1935–1968.
- Kroll, L. E., & Lampert, T. (2011). Unemployment, social support and health problems: Results of the GEDA study in Germany, 2009. *Deutsches Ärzteblatt International*, 108.
- Lain, D. (2016). *Reconstructing Retirement: Work and Welfare in the UK and USA*. Bristol: Policy press.
- Lynn, P. (2009). Sample design for understanding society. *Underst. Soc. Work. Pap. Ser* 2009.
- MacLeod, P., Fitzpatrick, A., Hamlyn, B., Jones, A., Kinver, A., & Page, L. (2012). Attitudes to Pensions: The 2012 survey. *Department of Work and Pension: Research Report*, 813.
- Marcus, J. (2013). The effect of unemployment on the mental health of spouses—Evidence from plant closures in Germany. *Journal of Health Economics*, 32(3), 546–558.
- Marmot, M. G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., ... Smith, G. D. (1991). Health inequalities among British civil servants: The Whitehall II study. *The Lancet*, 337(8754), 1387–1393.
- Mazzonna, F., & Peracchi, F. (2017). Unhealthy retirement? *Journal of Human Resources*, 52(1), 128–151.
- Mitra, S., & Jones, K. (2017). The impact of recent mental health changes on employment: New evidence from longitudinal data. *Applied Economics*, 49(1), 96–109.
- Moon, J. R., Glymour, M. M., Subramanian, S., Avendaño, M., & Kawachi, I. (2012). Transition to retirement and risk of cardiovascular disease: Prospective analysis of the US health and retirement study. *Social Science & Medicine*, 75(3), 526–530.
- Munnell, A., Sanzenbacher, G., & Rutledge, M. (2015). What causes workers to retire before they plan?, Centre For Retirement Research at Boston College.
- OECD. (2013). *Pensions at a Glance 2013: OECD and G20 Indicators*. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2015). *Education at a Glance 2015*.
- OECD. (2016). "Pensions at a Glance (Edition 2016)", OECD Pensions Statistics (database), in (Editor Ed.) Eds.). In *Book "Pensions at a Glance (Edition 2016)", OECD Pensions Statistics (database)*. City: Paris.
- OECD. (2017a). *Pensions at a glance 2017: OECD and G20 indicators*. Paris: OECD Publishing.
- OECD. (2017b). *Preventing ageing unequally*. Paris: OECD publishing.
- OECD/EU. (2018). *Health at a Glance: Europe 2018*.
- Paccagnella, M. (2016). *Age, ageing and skills*. Paris: OECD publishing.
- PPI. (2015). *The pension primer: A guide to the UK pension system*. London: Pension Policy Institute.
- Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M. R., & Rahman, A. (2007). No health without mental health. *The Lancet*, 370(9590), 859–877.
- Ravesteijn, B., Kippersluis, H. V., & Doorslaer, E. V. (2018). The wear and tear on health: What is the role of occupation? *Health Economics*, 27(2), e69–e86.
- Rechel, B., Grundy, E., Robine, J.-M., Cylus, J., Mackenbach, J. P., Knai, C., & McKee, M. (2013). Ageing in the European Union. *The Lancet*, 381(9874), 1312–1322.
- Santi, I., Kroll, L. E., Dietz, A., Becher, H., & Ramroth, H. (2013). Occupation and educational inequalities in laryngeal cancer: The use of a job index. *BMC Public Health*, 13(1), 1080. <https://doi.org/10.1186/1471-2458-13-1080>
- Schmitt, J. S., & Di Fabio, R. P. (2004). Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *Journal of Clinical Epidemiology*, 57(10), 1008–1018.
- Schmitz, H. (2011). Why are the unemployed in worse health? The Causal Effect of Unemployment on health. *Labour Economics*, 18(1), 71–78.
- Shai, O. (2018). 'Is retirement good for men's health? Evidence using a change in the retirement age in Israel. *Journal of Health Economics*, 57, 15–30.
- Sobotka, T., & Toulemon, L. (2008). Overview Chapter 4: Changing family and partnership behaviour: Common trends and persistent diversity across Europe.
- Solovieva, S., Pensola, T., Kausto, J., Shiri, R., Heliövaara, M., Burdorf, A., ... Viikari-Juntura, E. (2014). Evaluation of the validity of job exposure matrix for psychosocial factors at work. *PLoS ONE*, 9(9), e108987.

- Staubli, S., & Zweimüller, J. (2013). Does raising the early retirement age increase employment of older workers? *Journal of Public Economics*, 108, 17–32.
- Thurley, D., & Keen, R. (2017). *State pension age increases for women born in the 1950s*. London: House of Commons Library.
- van Solinge, H., & Henkens, K. (2017). Older workers' emotional reactions to rising retirement age: The case of the Netherlands. *Work, Aging and Retirement*, 3(3), 273–283.
- Ware, J. E. (2002). *The SF-12v2TM how to score version 2 of the SF-12® health survey: (with a supplement documenting version 1)*: Quality metric.
- Webber, D., & Mallet, L. (2017). What has happened to the income of retired households in the UK over the past 40 years?, in (Editor Ed.) ^ Eds.). In *Book What has happened to the income of retired households in the UK over the past 40 years?* City: ONS Office of National Statistics. Newport.
- Wester, G., & Wolff, J. (2010). The social gradient in health: How fair retirement could make a difference. *Public Health Ethics*, 3(3), 272–281.
- WHO. (2013). 'Mental health action plan 2013–2020', in (Editor Ed.) Eds.). In *Book Mental health action plan 2013–2020*. City: World Health Organization.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Carrino L, Glaser K, Avendano M. Later retirement, job strain, and health: Evidence from the new State Pension age in the United Kingdom. *Health Economics*. 2020;29:891–912. <https://doi.org/10.1002/hec.4025>